

高品質データベース構築  
と個人情報を含むデータの利活  
用における課題  
—大阪大学における取り組みを例に—

八木康史

大阪大学 理事・副学長

データビリティフロンティア機構長

# データバリティサイエンス構想



大阪大学  
OSAKA UNIVERSITY

OPEN 2021

様々な学問分野を通じた  
データ収集、分析、活用の重要性

そのためのデータ利活用の高度化  
Datability Science

データバリティフロンティア機構  
(平成28年4月に設置)

大阪大学にある様々な関連部門の  
連携、強化による学際研究推進

Datability is all about the ability to use large volumes of data **sustainably** and **responsibly**.

[CeBit 2014, held in Hannover, Germany]

データバリティとは、大規模なデータを持続可能かつ責任ある形で活用する能力のことです。

[CeBit 2014 (ドイツ、ハノーバー)にて提唱]

## 【sustainable】

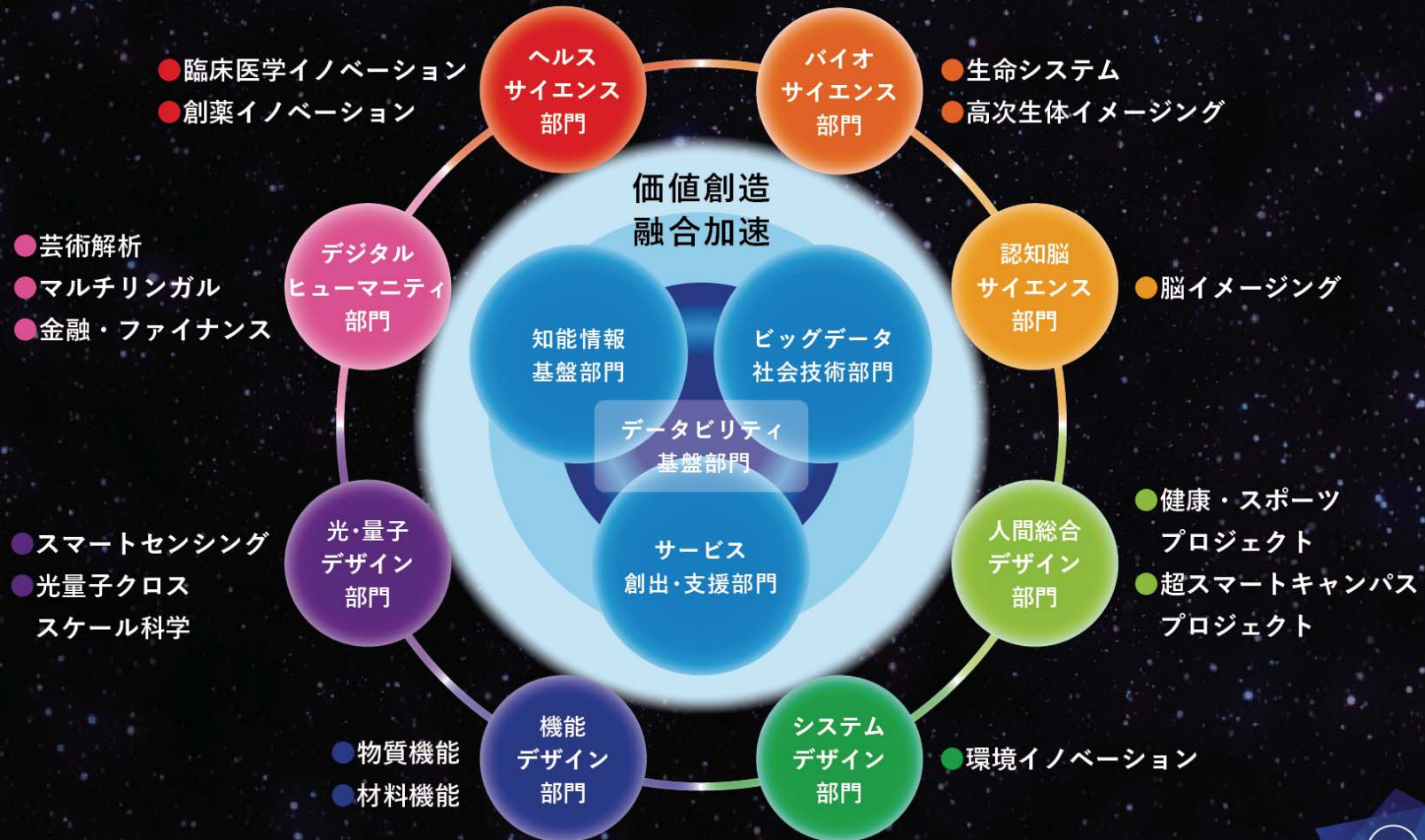
- ・継続的な対応を可能にするリソース整備
- ・データ・マネジメント&ガバナンスの確立
- ・セキュリティ対策

## 【responsible】

- ・社会問題／環境問題の解決(スマートシティ、ヘルスケア、エネルギー活用、経済活動等)
- ・プライバシー／個人情報の問題

データバリティサイエンティストの養成  
データバリティエンジニアの養成

# データビリティフロンティア機構



# 「データビリティ」におけるミッション



大阪大学  
OSAKA UNIVERSITY

OPEN 2021

## 1 データ駆動型研究推進

- 研究者マッチング
- 分野固有のノウハウを考慮した研究計画の設計

## 2 2次利用可能なデータベースと世界標準

- 大学という公共性を利用し、オプトイン、匿名化処理による被験者データ収集

- ◆ 大阪大学キャンパスを超スマートシティと想定した、1000人エクストラによる群衆行動データ、教職員ボランティアによるライフログデータ
- ◆ 生物学的・医学的データと臨床データの連結データベース

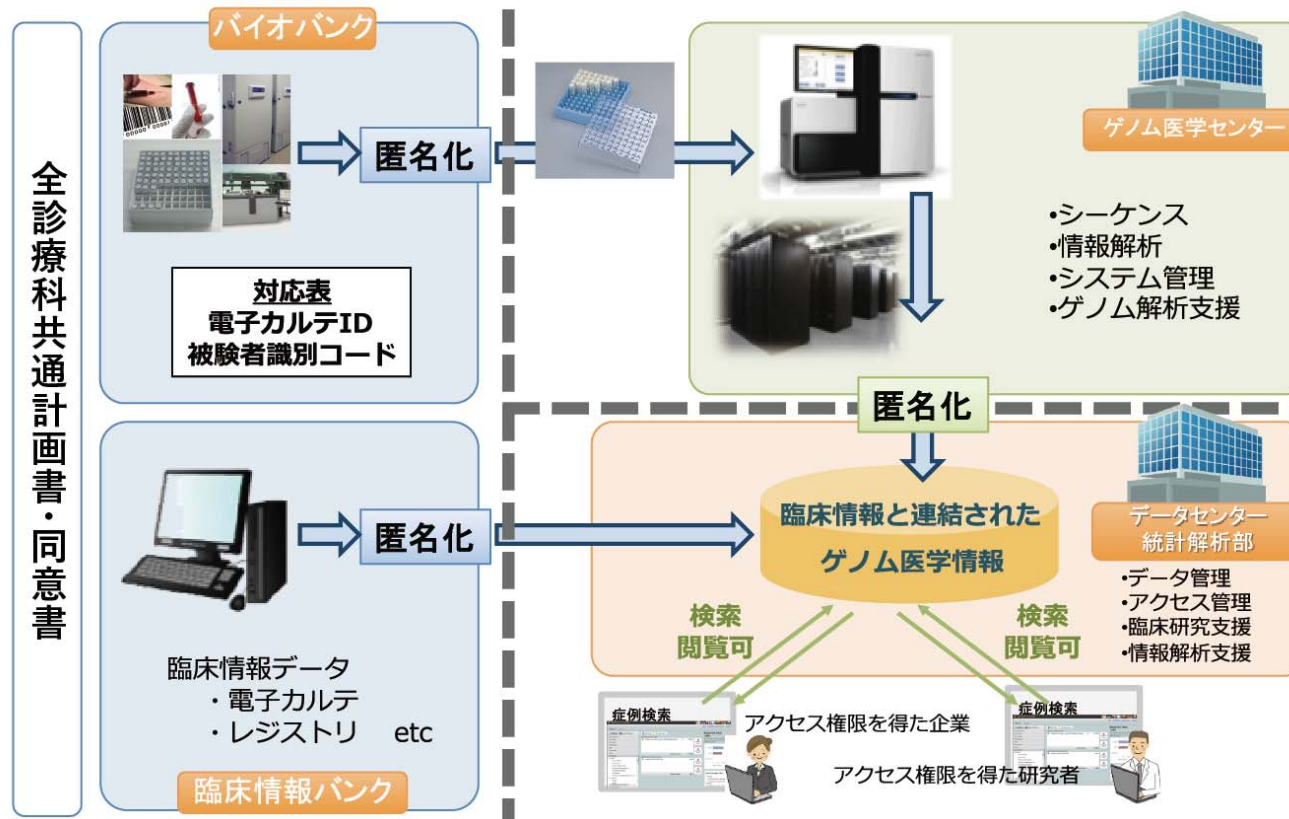


## 3 実践型人材育成

- OJTによる課題解決実践プログラム
- 基礎から実践までの一貫教育（講義とOJTの併用）
- ニーズに合わせたアラカルト形式の教育プログラム

# 医療イノベーションとデータビリティ

## ◆事例：連結データベースの構築（データマッチング）



大阪大学における診療科・関連病院の垣根を超えた医療データ集積

# —大阪大学の現状—

## 多数の大規模関連施設を有する臨床研究中核病院である医学部附属病院

### 独自の技術基盤

- 電子カルテの記録を入力、テンプレートで構造化
- 必要データを自動的に抽出してデータセンターに集める方法で、多施設共同臨床研究の実績

### 関連病院が集積

- 大阪府内には大きな病院が集結していて、連携しやすい環境である。これは他府県にはあまり見られない。
- 臨床研究中核病院として1万ベッドのネットワーク化がスタート。



# 厚生労働省の「医療介護分野の個人情報保護ガイドライン」における匿名化

当該個人情報から、当該情報に含まれる氏名、生年月日、住所等、個人を識別する情報を取り除くことで、特定の個人を識別できないようにすることをいう。顔写真については、一般的には目の部分にマスキングすることで特定の個人を識別できないと考えられる。なお、必要な場合には、その人と関わりのない符号又は番号を付すこともある。このような処理を行っても、事業者内で医療・介護関係個人情報を利用する場合は、事業者内で得られる他の情報や匿名化に際して付された符号又は番号と個人情報との対応表等と照合することで特定の患者・利用者等が識別されることも考えられる。法においては、「他の情報と容易に照合することができ、それにより特定の個人を識別することができることとなるもの」についても個人情報に含まれるものとされており、

匿名化に当たっては、当該情報の利用目的や利用者等を勘案した処理を行う必要があり、あわせて、**本人の同意を得るなどの対応も考慮する必要がある**。また、特定の患者・利用者の症例や事例を学会で発表したり、学会誌で報告したりする場合等は、氏名、生年月日、住所等を消去することで匿名化されることが考えられるが、**症例や事例により十分な匿名化が困難な場合は、本人の同意を得なければならない**。

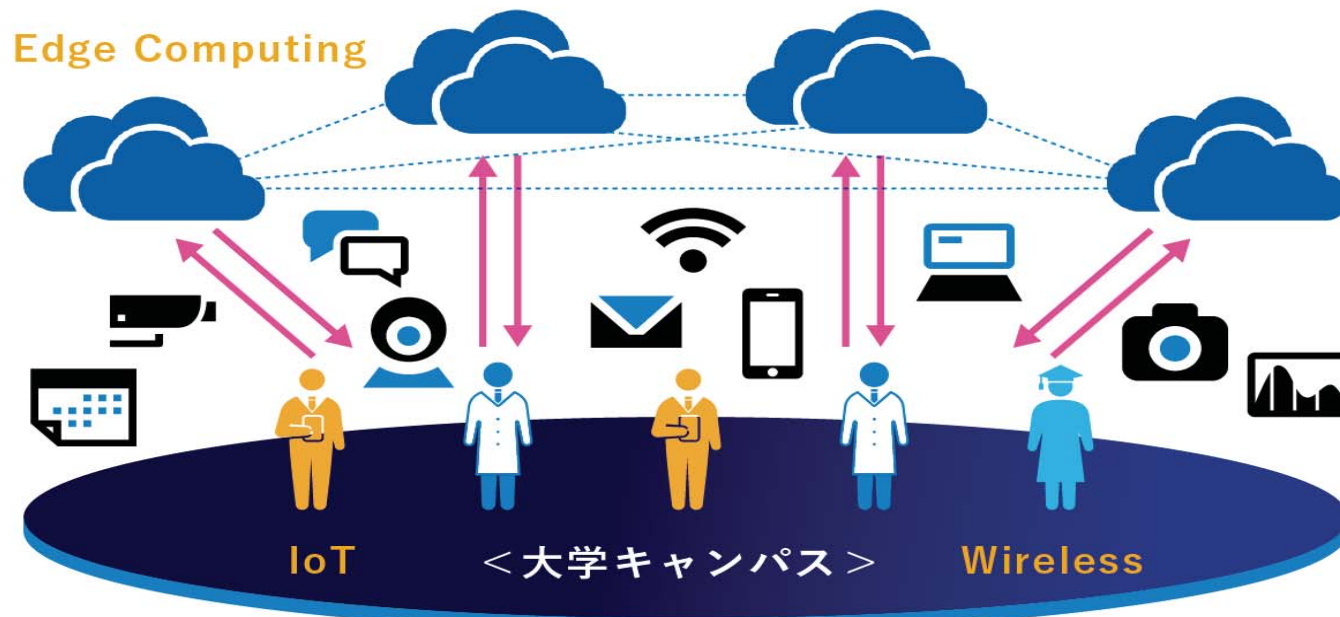
# 臨床データの利用価値

- AMED 平成28年度「臨床ゲノム情報統合データベース整備事業」に係る公募について
  - (a) 希少・難治性疾患領域における臨床ゲノム情報データストレージの整備に関する研究
  - (b) がん領域における臨床ゲノム情報データストレージの整備に関する研究
  - (c) 感染症領域における臨床ゲノム情報データストレージの整備に関する研究
  - (d) 認知症及びその他領域における臨床ゲノム情報データストレージの整備に関する研究
- データはただ取るだけでは、意味がない。目的を定めた中でのデータベース化が必要
- また、日常の生活データとの連結、脳科学との連結は、新たな研究を生み出す。その際に健康者データも重要



# 超スマートキャンパスプロジェクト

## <スマートキャンパス基盤>



- ◆大阪大学内の人々の活動状況センシング+キャンパス内ライフログ収集
- ◆アンビエントな情報サービスの提供
  - 授業・講演会、実験、通学、防犯に役立つサービス、訪問者支援
  - 研究室や大学のエネルギー削減
  - 教職員 & 学生の健康管理や見守り、クラブ活動の安全性確保など
- ◆大学内のシーズ技術を活用した知識創出

# 健康・スポーツプロジェクト

## スポーツ医科学の ニーズ

- 離れた“選手”の状態を知りたい  
→離れた場所の状態を知る
- “選手”の多くの情報、多くの選手の状態を知りたい  
→多くの状態を知る
- 次の局面へ対応したい  
→今の状態、近い将来の状態を知る

## 情報科学の 技術シーズ

- 無線技術
- ビッグデータ技術
- リアルタイム  
処理・予測技術

- ◆競技力向上 (HP: High Performance)
- ◆障害予防 (IP: Injury Prevention)
- ◆競技復帰 (RTP: Return to Play)

生体  
情報



外部  
情報

- ・バイタルセンサー (心拍、呼吸、体温、汗など)
- ・環境センサー (気温、湿度、風、気圧、WBGTなど)
- ・位置センサー (GPS、レーザーなど)
- ・モーションセンサー (加速度、傾斜、衝撃など)
- ・パフォーマンスセンサー (球速、ラケットスピードなど)
- ・力学センサー (足圧、握力など)



インターネット  
ゲートウェイ

# 大阪大学から大都市「うめきた」へ



OPEN 2021

240万人／日が往来する大阪駅の北側での「うめきた2期区域開発地域」

## 産学官民による共創イノベーション



### ライフデザイン・イノベーション

超スマート社会が到来する中、IoTやビッグデータ等の活用により、創薬や医療などの分野にとどまらず人々が健康で豊かに生きるための新しい製品・サービスを創出する。

# 利用目的の特定、公開ならびに第三者提供

## 【第15条】

個人情報取扱事業者は、個人情報を取り扱うに当たっては、その利用の目的をできる限り特定しなければならない。

## 【第15条第2項】

個人情報取扱事業者は、利用目的を変更する場合には、変更前の利用目的と相当の関連性を有すると合理的に認められる範囲を超えて行ってはならない。

## 【第18条第3項】

個人情報取扱事業者は、利用目的を変更した場合は、変更された利用目的について、本人に通知し、又は公表しなければならない。

## 【第23条】

個人情報取扱事業者は、次に掲げる場合を除くほか、あらかじめ本人の同意を得ないで、個人データを第三者に提供してはならない。

# 人のデータを扱う上で

- 利用目的を明確に定めたデータ収集
- オプトインによるデータ収集
  
- 上記を満たした上での高付加価値データは極めて競争力あり
- 高付加価値とは、連結データ化に加え、
- データに対するアノテーション
  - 例) CT画像中でのがん領域のセグメンテーション結果、所見
  - 例) 監視映像における人物領域と追跡結果、人物属性 (性別、年齢等)